13 Lathers C, Schraeder P, Autonomic dysfunction in epilepsy: characterisation of autonomic cardiac neural discharge associated with pentylenetetrazol-induced epileptogenic activity. *Epilepsia* 1982;**23**:633-47
14 Lathers C, Schraeder P, Weiner F. Synchronization of cardiac autonomic neural discharge with epileptogenic activity: the lockstep phenomenon. *Electroencephalogr Clin Neurophysiol* 1987;**67**:247-59
15 Dasheiff R, Dickinson L. Sudden unexpected death of epileptic patient due to cardiac arrhythmia after seizure. *Arch Neurol* 1986;**43**:194-6
16 Mufson E, Mesulam M. Insula of the old world monkey. II: Afferent cortical input and comments on the claustrum. *J Comp Neurol* 1982;**212**:23-37
17 Cechetto D, Saper C. Role of the cerebral cortex in autonomic function. In: Loewy A, Spyer K, eds. *The autonomic nervous system: central regulation of autonomic functions.* Oxford: Oxford University Press, 1988
18 Oppenheimer SM, Hachinski VC, Cechetto DF. Cardiac chronotopic organization of the rat insula. *Neurosci Abstracts* 1989;**15**:595

Doubt and certainty in statistics

How can our beliefs be verified? The Continental philosophers, Descartes, Leibniz, Kant and Hegel held that logic alone would establish the truth; by contrast, the British empiricists, Bacon, Locke, Hume, Mill and the American pragmatists, Pierce, James, and Dewey, attached more importance to observation. The Europeans were influenced by the spectacular successes of Greek mathematics; the equally splendid achievements of science guided the Anglo-US philosophers. Both are right in the relevant context and each has a bearing on statistical theory.

A theorem in mathematics starts and ends in the mind. Given the initial premisses only logic is needed to reach the final answer. But problems arise when the argument starts, not from axioms, but from sense data of the real world. More than one theory will account for the observations and logic may not, by itself, settle the question. After centuries of theoretical argument the solution emerged in the 17th century. A crucial experiment could show which of two contradictory ideas was to be preferred. Aristotelian theory proposed that acceleration in falling bodies was proportional to mass; Galileo's experiments showed it to be constant. For 1500 years Galen's explanation of the pulse was that blood ebbed and flowed; Harvey's experiments showed that it circulated. In philosophic terms, deductive argument from theory gave way to inductive argument from data - but Harvey's own words have a more Shakespearian ring: 'I profess both to learn and to teach, not from the positions of philosophers, but from the fabric of Nature'.

Since that time immense efforts have been made to codify the scientific method, to give the inductive argument the same kind of logical elegance as mathematics. This is now widely recognized as a hopeless endeavour. External reality on the one hand and the images we produce in our consciousness, on the other, are not the same. They differ vastly. Can the rules of logic developed in one be expected to apply rigorously in the other? Of course the question does not arise in pure mathematics, but it is never far from our thoughts in science.

A school of philosophy, which was once widely supported, maintained that the external world exists only in our minds. If that is rejected then it follows that all scientific arguments, however excellent, in the final analysis use mental images to represent objects that are not mental. An element of approximation is unavoidable; light appears to be a continuous wave and also discrete particles. In the end a scientific theory is accepted not because it is 'true', whatever that may mean, but because it works and is useful. Some helpful rules have emerged. The prime test of a theory is that it should predict correctly. Secondly it must be consistent with the rest of science. It must have, as Einstein[1] put it, both 'internal and external coherence'. A crucial experiment never verifies the 'correct' idea in any absolute sense; Popper[2] argues that its most valuable function is to falsify the wrong idea. The successful theory will itself be discarded if new facts come to light which rule it out. Thus science advances by making progressively better approximations to reality. Pythagoras is constant for eternity; Science is an asymptotic variable.

In the past, statistics was regarded as mathematics rather than science. Certainly pure mathematics has always played a crucial role. Pascal, Fermat and Bernouilli applied the techniques developed in pure mathematics to predicting the behaviour of dice. That was the start and Neyman[3] shows that later developments have followed a similar route. However, mathematics has not succeeded in formulating a fundamental theory of probability. Fisher[4,5] wrestled with this problem for more than 10 years. The distinguished mathematician Littlewood[6] concludes that neither the frequency theory of probability, nor any of the others, is satisfactory in a strictly mathematical sense; they cannot be rigorously proved in the way that Euclid's theorems are proved.

Should we therefore use the modern computer to test statistics by experiment? Littlewood rejects experimental proof because false theories can give correct predictions. This objection is valid in pure mathematics when logic provides a more conclusive proof than empirical evidence. But in statistics, if mathematical proof fails us, surely experimental proof is the next best option. It may not be possible to justify the scientific method by formal logic, but we owe to it everything that distinguishes our world from the Dark Ages. If a satisfactory mathematical theory of probability is ever created then it will prevail. But meanwhile there seems a strong case for regarding statistics not so much as mathematics, but more as an exact science, akin to physics. The ultimate court of appeal is not the reductio ad absurdum, but the crucial experiment.

There is a second, more fundamental, reason why statistics should be subject to the same discipline

0141-0768/90/030136-02/$02.00/0

as science. Suppose, for the sake of argument, that a satisfactory mathematical theory of probability were produced could we then say that it must apply in the real world? In the light of modern physics, that seems uncertain. Einstein considered it profoundly mysterious that ideas created wholly in the mind should have any relevance at all to external reality - yet Euclidian geometry works very well on the terrestrial scale. But it fails on the astronomical scale where the geometry of Riemann fits better. 'As far as the propositions of mathematics refer to reality they are not certain; as far as they are certain they do not refer to reality.'[7] It follows that even if a statistical theory were mathematically perfect it would not necessarily give the right answer in real life - only observation can settle that.

Certainly statistics merits study for its own sake - the grandeur of the central limit theorem alone would justify that. But it is not only pure mathematics. It is also needed in real life and these areas must be answerable to the same rules as science. The risk with mathematics is that as ideas become more abstract they may at some point detach from reality, leading to decades of theoretical argument. It is then a useful discipline to translate the mathematics into words and visual images and, if necessary, into a computer simulation. Abstract ideas come more sharply into focus when translated into real terms. Any significance test which fails this challenge should, in our view, be discarded. Pure mathematics need not apply to real life, but statistics does have that duty. These conclusions apply to Fisher's Exact Test, for example[8].

The sceptical doctor can easily reassure himself that the significance tests he commonly uses are verifiable - experimental P values, from computer simulations, confirm the predictions. But there are debateable areas. Two examples illustrate the point, one simple, the other complex:

(1) Some statisticians, arguing from the method of maximum likelihood, conclude that the standard variance formula should have n as denominator, not $n-1$. But experiment shows that only $n-1$ gives the right answer. All that is needed is an inexpensive home computer and a simple programme of six lines.

If the variance is found, not from a sample, but from the whole population then n is right. Or, if the formula for computing the variance uses, not the sample mean, but the true mean for the whole population then again n is right. These potentially confusing assertions become clear with the corresponding simulations.

(2) Finding the curve of best fit for non-linear functions by the method of least squares, using Taylor's Theorem, surely ranks as a triumph of pure mathematics. It works extremely well for high quality data. But it may fail if the variance is rather large. The theory assumes that if the data are Gaussian then the derived parameters will likewise follow a symmetric distribution. Computer simulations show this to be true in general, but as the variance of the simulated data is increased so the results may begin to skew and the mean values to deviate from what is predicted. In such cases the simulation is valuable in three ways: it shows which mathematical assumption is unsound. It shows how much larger n must be to get it right. If n cannot be increased the simulation itself enables a reasonably accurate correction to be applied. In passing it provides an elegant demonstration of the central limit theorem.

The method can be adapted to any non-linear function by differentiating the basic equation and Barlow[9] shows that it can be safely used by those who lack the benefit of expert statistical help. Nevertheless there is much to be said for cross-checking the mathematics against a computer simulation. The simulations needed to check complex equations are surprisingly simple. One caveat is needed. An efficient random number generator is essential and those provided on some computers are not adequate; we advocate that of Wichmann & Hill[10].

Computer simulations can sometimes solve problems which resist mathematical analysis. The experimental P values are generally correct to two figures, which is more accurate than some approximations commonly used in significance testing. Teaching statistics will be revolutionized, especially for those who do not find mathematics easy - abstract ideas become more concrete. In the past statistics required long exercises in elementary algebra; computer programmes make this superfluous, leaving doctors free to think about the underlying logic. This should reduce the number of howlers published.

Lastly the aesthetic factor is important. W H Auden said he could not define poetry but he recognized it from his symptoms. We suggest that the effect results from the landscape we know resonating, as it were, with the image created by the artist. In the same way a computer simulation may seem banal in itself, but is transfigured by the underlying equation - each illuminates the other. Even so dedicated a pure mathematician as G H Hardy[11] concedes that the beauty of a theorem may be enhanced if it is seen in relation to the real world. 'One thing I have learned in a long life: that all our science, measured against reality, is primitive and child-like - and yet it is the most precious thing we have.'[12]

R S Cormack
Northwick Park Hospital & Clinical Research Centre, Harrow
Nathan Mantel
The American University, 4900 Auburn Avenue, Bethesda, USA

References

1 Einstein AE. Einstein, science and culture. In: Finch AP, ed. *Centenary volume*. London: Heinemann, 1979:168
2 Popper K. Science: conjectures and refutations. *Conjectures and refutations*. London: Routledge, 1963
3 Neyman J. In: Owen DB, ed. *The history of statistics and probability*. New York: Decker, 1976:149-50
4 Fisher RA. The logic of inductive inference. *J R Statist Soc (A)* 1935;**98**:39-54
5 Fisher RA. The logical inversion of the notion of a random variable. *Sankhya* 1945;**7**:129-32
6 Littlewood JE. *Littlewood's miscellany*. Cambridge: Cambridge University Press, 1986:71-3
7 Einstein AE. Approaches to the teaching of relativity. In: Finch AP, ed. *Centenary volume*. London: Heinemann, 1979:252
8 Cormack RS. The meaning of probability in relation to Fisher's Exact Test. *Metron* 1986;**44**:1-30
9 Barlow RB. Line-fitting by least squares. In: *Biodata handling with microcomputers*. Cambridge: Elsevier, 1983:94-9
10 Wichmann BA, Hill ID. An efficient, portable pseudo-random number generator. *J R Statist Soc (C)* 1982; **31**:188-90
11 Hardy GH. *A mathematician's apology*. Cambridge. Cambridge University Press, 1969:85-90,127.
12 Einstein AE. Einstein and education. In: Finch AP, ed. *Centenary volume*, London: Heinemann, 1979:218